



Draft Genome Sequences of Three *Shewanella* sp. Strains Isolated from Urban Freshwaters in the Ohio River Valley, USA

Joy L. Kappesser,^a  Joshua T. Cooper^a

^aDepartment of Biological Sciences, Northern Kentucky University, Highland Heights, Kentucky, USA

ABSTRACT Draft genome sequences of three *Shewanella* sp. strains are reported. The strains NKUCC01_JLK, NKUCC05_KAH, and NKUCC06_TVS were isolated from freshwater sources in the Ohio River Valley, USA. These genome sequences provide insights into *Shewanella* adaptation to urban freshwaters and may help to elucidate their roles in biogeochemical cycling.

Shewanella species are facultative anaerobic, Gram-negative gammaproteobacteria that are primarily aquatic and inhabit diverse environments (1). The diversity of *Shewanella* species is likely underestimated (2), with many uncultivated representatives performing unknown roles in ecological processes. By studying the genomes of *Shewanella* species from urban freshwaters, we hope to evaluate and infer their roles in freshwater biogeochemical cycling.

Shewanella sp. strain NKUCC01_JLK and strain NKUCC05_KAH were isolated from subsurface water grab samples collected with a sterile Whirlpak bag from the Licking River, Kentucky, at Frederick's Landing in January 2021, while strain NKUCC06_TVS was isolated from subsurface water grab samples collected with another sterile Whirlpak bag from Loch Norse on the campus of Northern Kentucky University in January 2021 (coordinates are available in Table 1). Water samples were serially diluted 1:10, spread onto tryptic soy agar (TSA), and grown for 48 h at 25°C. Individual colonies were successively purified on TSA for several months using quadrant streaking. For each strain, genomic DNA was isolated from cells grown in tryptic soy broth at 25°C and was extracted using the UltraClean microbial DNA isolation kit (Qiagen, MD, USA). DNA concentrations were quantified using the Qubit 3.0 broad range kit (Invitrogen) before samples were sent to the Microbial Genome Sequencing Center (MiGS) at the University of Pittsburgh (Pittsburgh, PA). The samples were sequenced to the requested depths using the Illumina NextSeq 2000 platform (2 × 150 bp) to generate paired-end reads. Briefly, libraries were prepared using Illumina Nextera chemistry kits, according to the manufacturer's protocol, with IDT Illumina 10-bp indices. Reads were examined with FastQC v0.11.5 (3), trimmed with fastp (4), and assembled with Shovill v1.1.0 (5). Assemblies were then annotated with the NCBI Prokaryotic Genome Annotation Pipeline (PGAP) v5.2 (6). Coverage was determined using the read aligner BMap v38.90 (7) and the trimmed reads. CheckM (8) was used to estimate genome completeness, and data were explored using Kbase v2.3.2 (9). Default software parameters were used throughout unless specified otherwise. Features of the three *Shewanella* sp. strains are summarized in Table 1.

To explore the taxonomy of our strains, we used Tetra Correlation Search with JSpeciesWS v3.8.2 (10) to select genomes sharing >97% identity. We supplemented the search using PATRIC v3.6.9 Similar Genome Finder (11) and added three closely related genomes from NCBI (*Shewanella putrefaciens* HRCR-6 [GenBank accession number [NZ_JAEU000000000](#) and RefSeq accession number [GCF_000519065.1](#)], *Shewanella putrefaciens* 97 [GenBank accession number [NZ_QNSC000000000](#) and RefSeq accession number [GCF_003315425.1](#)], and *Shewanella* sp. strain WE21 [GenBank accession number [NZ_CP023019](#) and RefSeq accession number [GCF_002966515.1](#)]). Pairwise Tetra Correlation analysis indicated that all strains had

Citation Kappesser JL, Cooper JT. 2021. Draft genome sequences of three *Shewanella* sp. strains isolated from urban freshwaters in the Ohio River Valley, USA. Microbiol Resour Announc 10:e00742-21. <https://doi.org/10.1128/MRA.00742-21>.

Editor J. Cameron Thrash, University of Southern California

Copyright © 2021 Kappesser and Cooper. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](#).

Address correspondence to Joshua T. Cooper, cooperjo@nku.edu.

Received 26 July 2021

Accepted 17 August 2021

Published 16 September 2021

TABLE 1 Summary of genome data for the three *Shewanella* sp. strains described in this study

Isolate	Collection site coordinates	Collection site habitat	Assembly size (bp)	G+C content (%)	No. of contigs	N_{50} (bp)	Coverage (x)	Size (bp)	Completeness (%)	Contamination (%)	No. of protein-coding sequences	No. of tRNA genes	No. of RNA genes	SRA accession no.	GenBank accession no.	BioSample accession no.
NKUCC01_JLK	39.050611°N, 84.493150°W	Fredericks Landing, Licking River, KY shoreline of	5,109,354	45.21	62	227,063	192	6,688,222	99.68	0.67	4,425	109	129	SRR15168786	JAHKRD0000000000	SAMIN19515366
NKUCC05_KAH	39.050611°N, 84.493150°W	Fredericks Landing, Licking River, KY shoreline of	5,057,459	45.15	87	176,748	179	6,185,428	99.84	0.41	4,335	100	133	SRR15168782	JAHKQZ0000000000	SAMIN19515369
NKUCC06_TV5	39.031240°N, 84.462720°W	Licking River, KY Loch Norse, Northern Kentucky University, KY	5,050,477	45.72	71	223,290	172	5,903,314	99.84	0.37	4,289	99	135	SRR15168780	JAHKQX0000000000	SAMIN19515371

values of >0.999 with respect to each other and the three additional genomes. Average nucleotide identity on BLAST+ (ANIb) values indicate that all strains and the three additional genomes likely belong to the same species, with ANIb values of >96.7%. Digital DNA-DNA hybridization (dDDH) from the Type (Strain) Genome Server (TYGS) v295 (12) with formula d4 indicated that all three strains shared >70% identity with each other but shared only 40% with any type *Shewanella* species.

In summary, the strains NKUCC01_JLK, NKUCC06_TVS, and NKUCC05_KAH may be representatives of an unnamed clade of freshwater *Shewanella* (with *S. putrefaciens* HRCR-6, *S. putrefaciens* 97, and *Shewanella* sp. strain WE21), as illustrated by Thorell et al. (13), which adds to our understanding of freshwater *Shewanella* diversity.

Data availability. The whole-genome shotgun projects for *Shewanella* sp. strain NKUCC01_JLK, strain NKUCC05_KAH, and strain NKUCC06_TVS have been deposited in DDBJ/ENA/GenBank under the BioProject accession number [PRJNA734631](https://doi.org/10.6084/m9.figshare.14999802); the versions described in this paper are versions [JAHKRD000000000.1](https://doi.org/10.6084/m9.figshare.14999802), [JAHKQZ000000000.1](https://doi.org/10.6084/m9.figshare.14999802), and [JAHKQX000000000.1](https://doi.org/10.6084/m9.figshare.14999802), respectively. Detailed information can be found in Table 1 regarding BioSample and Sequence Read Archive (SRA) accession numbers. Log files corresponding to the output of the genome assembly by Shovill v1.1.0 and the subsequent graph assembly files for each strain (.gfa) are located at FigShare (<https://doi.org/10.6084/m9.figshare.14999802>).

ACKNOWLEDGMENTS

This research was supported by the Northern Kentucky University Department of Biological Sciences, and the high-throughput computing resources were made possible by the National Institutes of Health (grant P20GM103436).

We thank Trevor Shumate and Katherine Holtz for isolation of strains NKUCC06_TVS and NKUCC05_KAH. We thank Kathleen Schenck for reviewing earlier versions of the manuscript, and we thank the two anonymous reviewers for their helpful comments and feedback.

REFERENCES

- Lemaire ON, Méjean V, Iobbi-Nivol C. 2020. The *Shewanella* genus: ubiquitous organisms sustaining and preserving aquatic ecosystems. *FEMS Microbiol Rev* 44:155–170. <https://doi.org/10.1093/femsre/fuz031>.
- Li B-B, Cheng Y-Y, Fan Y-Y, Liu D-F, Fang C-Y, Wu C, Li W-W, Yang Z-C, Yu H-Q. 2018. Estimates of abundance and diversity of *Shewanella* genus in natural and engineered aqueous environments with newly designed primers. *Sci Total Environ* 637–638:926–933. <https://doi.org/10.1016/j.scitotenv.2018.05.051>.
- Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>.
- Chen S, Zhou Y, Chen Y, Gu J. 2018. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34:i884–i890. <https://doi.org/10.1093/bioinformatics/bty560>.
- Seeman T. 2007. Shovill v1.1.0. <https://github.com/tseemann/shovill>.
- Tatusova T, DiCuccio M, Badretdin A, Chetvernin V, Nawrocki EP, Zaslavsky L, Lomsadze A, Pruitt KD, Borodovsky M, Ostell J. 2016. NCBI Prokaryotic Genome Annotation Pipeline. *Nucleic Acids Res* 44:6614–6624. <https://doi.org/10.1093/nar/gkw569>.
- Bushnell B. 2014. BBMap short read aligner, and other bioinformatic tools. <https://sourceforge.net/projects/bbmap>.
- Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. 2015. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. *Genome Res* 25:1043–1055. <https://doi.org/10.1101/gr.186072.114>.
- Arkin AP, Cottingham RW, Henry CS, Harris NL, Stevens RL, Maslov S, Dehal P, Ware D, Perez F, Canon S, Sneddon MW, Henderson ML, Riehl WJ, Murphy-Olson D, Chan SY, Kamimura RT, Kumari S, Drake MM, Brettin TS, Glass EM, Chivian D, Gunter D, Weston DJ, Allen BH, Baumohl J, Best AA, Bowen B, Brenner SE, Bun CC, Chandonia J-M, Chia J-M, Colasanti R, Conrad N, Davis JJ, Davison BH, DeJongh M, Devoid S, Dietrich E, Dubchak I, Edirisinghe JN, Fang G, Faria JP, Frybarger PM, Gerlach W, Gerstein M, Greiner A, Gurtowski J, Haun HL, He F, Jain R, Joachimiak MP, Keegan KP, Kondo S, Kumar V, Land ML, Meyer F, Mills M, Novichkov PS, Oh T, Olsen GJ, Olson R, Parrello B, Pasternak S, Pearson E, Poon SS, Price GA, Ramakrishnan S, Ranjan P, Ronald PC, Schatz MC, Seaver SMD, Shukla M, Sutormin RA, Syed MH, Thomason J, Tintle NL, Wang D, Xia F, Yoo H, Yoo S, Yu D. 2018. KBase: the United States Department of Energy Systems Biology Knowledgebase. *Nat Biotechnol* 36:566–569. <https://doi.org/10.1038/nbt.4163>.
- Richter M, Rosselló-Móra R, Oliver Glöckner F, Peplies J. 2016. JSpeciesWS: a web server for prokaryotic species circumscription based on pairwise genome comparison. *Bioinformatics* 32:929–931. <https://doi.org/10.1093/bioinformatics/btv681>.
- Davis JJ, Wattam AR, Aziz RK, Brettin T, Butler R, Butler RM, Chlenski P, Conrad N, Dickerman A, Dietrich EM, Gabbard JL, Gerdes S, Guard A, Kenyon RW, Machi D, Mao C, Murphy-Olson D, Nguyen M, Nordberg EK, Olsen GJ, Olson RD, Overbeek JC, Overbeek R, Parrello B, Pusch GD, Shukla M, Thomas C, VanOeffelen M, Vonstein V, Warren AS, Xia F, Xie D, Yoo H, Stevens R. 2020. The PATRIC Bioinformatics Resource Center: expanding data and analysis capabilities. *Nucleic Acids Res* 48:D606–D612. <https://doi.org/10.1093/nar/gkz943>.
- Meier-Kolthoff JP, Göker M. 2019. TYGS is an automated high-throughput platform for state-of-the-art genome-based taxonomy. *Nat Commun* 10: 2182. <https://doi.org/10.1038/s41467-019-10210-3>.
- Thorell K, Meier-Kolthoff JP, Sjöling Å, Martín-Rodríguez AJ. 2019. Whole-genome sequencing redefines *Shewanella* taxonomy. *Front Microbiol* 10: 1861. <https://doi.org/10.3389/fmicb.2019.01861>.